

Optimal Beamforming in Interference Networks with Perfect Local Channel Information

Rami Mochaourab, *Student Member, IEEE*, and Eduard Jorswieck, *Senior Member, IEEE*

Abstract

We consider settings in which T multi-antenna transmitters and K single-antenna receivers concurrently utilize the available communication resources. Each transmitter sends useful information only to its intended receivers and can degrade the performance of unintended systems. Here, we assume the performance measures associated with each receiver are monotonic with the received power gains. In general, the systems' joint operation is desired to be Pareto optimal. However, designing Pareto optimal resource allocation schemes is known to be difficult. In order to reduce the complexity of achieving efficient operating points, we show that it is sufficient to consider rank-1 transmit covariance matrices and propose a framework for determining the efficient beamforming vectors. These beamforming vectors are thereby also parameterized by $T(K - 1)$ real-valued parameters each between zero and one. The framework is based on analyzing each transmitter's power gain-region which is composed of all jointly achievable power gains at the receivers. The efficient beamforming vectors are on a specific boundary section of the power gain-region, and in certain scenarios it is shown that it is necessary to perform additional power allocation on the beamforming vectors. Two examples which include broadcast and multicast data as well as a cognitive radio application scenario illustrate the results.

Part of this work has been performed in the framework of the European research project SAPHYRE, which is partly funded by the European Union under its FP7 ICT Objective 1.1 - The Network of the Future. This work is also supported in part by the Deutsche Forschungsgemeinschaft (DFG) under grant Jo 801/4-1.

The authors are with the Department of Electrical Engineering and Information Technology, Dresden University of Technology, 01062 Dresden, Germany. E-mail: {Rami.Mochaourab,Eduard.Jorswieck}@tu-dresden.de. Phone: +49-351-46332239. Fax: +49-351-46337236.

Part of this work has been presented at IEEE International Zurich Seminar on Communications, Zurich, Switzerland, March 3–5, 2010 [1] and International Workshop on Cognitive Information Processing, Elba Island, Italy, June 14–17, 2010 [2].

I. INTRODUCTION

Interference is known to be one of the major factors that limits the performance of a communication system in a wireless network. This situation is common in multiuser settings when the systems concurrently share the available communication resources. In general interference networks, the performance measure of individual users is described by a utility function. This function depends in a monotonic way on the received signal power, interference signal power and noise power. The joint operation of the systems is efficient if it is not possible to improve the performance of one system without degrading the performance of another. In this case, the operating point is said to be Pareto optimal. It is always desired to design resource allocation schemes that lead to Pareto optimal operation points. In this way, the available communication resources are utilized efficiently to grant efficient operation of the systems. However, developing efficient resource allocation schemes is not straightforward and proves to be difficult. For instance, the problem of finding the maximum sum-rate or the proportional-fair operating point in the multiple-input single-output (MISO) interference channel (IC) is proven to be strongly NP-hard¹ [3].

In a multiuser setting, efficient operation of the systems requires the transmitters to maximize the power gain at intended receivers and also minimize the power gain at unintended receivers. In this work, we characterize the transmission strategies of each transmitter that are relevant to achieve Pareto optimal operating points. Moreover, we parameterize these by real values between zero and one. In this way, the set of efficient transmission strategies is confined and represented by low dimensional real parameters. This result tremendously reduces the complexity of designing efficient resource allocation schemes, and the parametrization can be utilized for low complexity coordination between transmitters.

We give a brief reference to related work in the MISO IC and a few of their applications. The MISO IC is an example of an interference network where the systems consist of transmitter-receiver pairs. For the two-user case, real-valued parametrization of each transmitter's efficient beamforming vectors is provided in [4]. The beamforming vectors that achieve Pareto optimal points are proven to be a linear combination of zero-forcing (ZF) transmission and maximum ratio transmission (MRT). Based on this characterization, a monotonic optimization framework is developed in [5] to find maximum sum-rate, proportional-fair and minimax operating points. The parametrization in [4] relates to a parametrization using the virtual SINR framework in [6]. The use of this framework is motivated by the design of distributed algorithms that require local channel state information (CSI) at each transmitter. This framework is extended to the precoding design in MIMO settings in [7]. The concept of combining the MRT and ZF strategies

¹Interestingly, these problems are efficiently solvable if rate requirements or interference constraints on each system are fixed.

is important for developing so-called distributed bargaining algorithms. These algorithms improve the operation of the systems from the noncooperative outcome [8], [9]. In [8], a distributed bargaining algorithm is developed which requires one bit signaling between the transmitters. Extension to the precoding design in the multiple-input multiple-output (MIMO) case is given in [10]. In [9], a similar distributed beamforming algorithm in the MISO IC is proposed for the case of statistical CSI at the transmitters. Also utilizing the parametrization in [4], a distributed bargaining process is proposed in [11] which requires four bit signaling between the transmitters. The process is proven to converge to an operating point arbitrarily close to the Pareto boundary and dominates the noncooperative outcome of the systems. In [12], the high signal to interference plus noise ratio (SINR) approximation of the achievable sum-rate of a system pair is utilized to determine suboptimal joint transmission strategies. The achieved performance is shown to be better than the joint MRT and joint ZF strategies.

In the K -user MISO IC, complex-valued parametrization of the Pareto boundary of the MISO IC rate-region is derived in [13] which requires $K(K - 1)$ complex-valued parameters in order to attain all Pareto optimal points. In [14], the K -user MISO IC is considered with the capabilities of time sharing the resources between the links. All points on the Pareto boundary of the MISO IC rate region are achieved with $K(K - 1)$ real valued parameters each between 0 and π . In [15], the authors characterize the Pareto boundary of the MISO IC through controlling interference temperature constraints (ITC) at the receivers. Each Pareto optimal rate tuple is achieved iteratively when each transmitter optimizes its transmission constrained by the ITCs. It is shown that $K(K - 1)$ real valued parameters, each between zero and a value depending on the channel vectors, are needed to achieve all Pareto optimal points. ITC is a terminology used in cognitive radio scenarios under the underlay paradigm [16]. It quantifies the amount of interference from the secondary transmitters that is tolerated by the primary users.

In [17], joint linear precoding is investigated taking into account the signaling overhead between the transmitters. The rate-region achieved with joint precoding is larger than the MISO IC rate-region, and all Pareto optimal beamforming vectors are parameterized by $K(K - 1)$ complex-valued parameters. For the same setting, a recent result in [18] reduces the number of parameters to $K + L$ real-valued scalars, each between zero and one, where L is the number of linear constraints on the transmission. Linear precoding MIMO IC algorithms are moreover investigated in [19] for a two-user system.

While the above mentioned results are provided for the MISO IC setting, the parametrization of efficient transmission strategies in a general multiuser setting is not straightforward. Moreover, neither the ITC-based [15] nor the Lagrangian-based [14], [18] characterizations can be generalized to our framework. A further example of a MISO multiuser setting which can be applied to our framework is when a

single transmitter sends common information to K single-antenna receivers. This setting corresponds to multicast transmission. Since the transmission rate depends on the weakest link in the system, the transmitter optimizes its transmission to achieve max-min-fairness at the receivers [20]. The multicast beamforming problem to achieve max-min-fairness is proven to be NP hard for $K \geq N$ [21], where N is the number of transmit antennas. In [22], the two-user multicast max-min-fair problem is studied, and the set of beamforming vectors which includes the solution of the max-min-fair problem is characterized.

In this work, we provide a general framework for parameterizing the transmission strategies of each transmitter which are relevant to achieve Pareto optimal points. This framework is applicable to settings where the utility functions of the systems are monotonic in the received power gains. The contributions and outline of this paper are as follows:

- We investigate the properties of efficient transmission of a single transmitter. These properties are acquired on studying the transmitter's power gain-region (Section III). The power gain-region is composed of all jointly achievable power gains at the receivers. Of interest are the transmission strategies which achieve its boundary part in a specific direction. We prove that the boundary of the power gain-region is convex and always achieved with single-stream beamforming (Lemma 3). Due to these properties, the corresponding strategies are characterized by real-valued parameters (Theorem 1). Furthermore, we characterize under which conditions power control is needed for efficient transmission. (i) When the number of transmit antennas is greater than or equal to the number of receivers K (Section III-A), we prove that full power transmission achieves all boundary points (Lemma 2). In this case, $K - 1$ real-valued parameters, each between zero and one, are needed to parameterize the beamforming vectors. (ii) When the number of transmit antennas is strictly less than the number of receivers (Section III-B), we characterize the transmission strategies for which power control is needed. For this case, an additional real-valued parameter between zero and one is needed that varies the power level at the transmitter.
- We utilize the developed single-transmitter framework for the multiple-transmitter case (Section IV). Based on the network setting and the monotonicity properties of each receiver's utility function, the boundary part which is relevant for Pareto optimal operation is determined for each transmitter's gain-region. Consequently, each transmitter's efficient strategies are parameterized (Theorem 2). We provide an example setting which includes broadcast and multicast data, and we apply the developed framework to this setting (Section IV-A). Moreover, we apply the framework to the K -user MISO IC (Section V-A). As a special case, the result for the two-user MISO IC in [13] follows. In addition,

we give an alternative characterization of the efficient transmission strategies (Corollary 1) which is motivated by the application of null-shaping constraints in underlay cognitive radio scenarios (Section V-B). We prove that all Pareto optimal operating points can be characterized through the design of null-shaping constraints on noncooperative secondary transmitters. Extensions to the case of multiple antennas at the receivers is covered in Section VI.

Notations: Column vectors and matrices are given in lowercase and uppercase boldface letters, respectively. The notation $x_{k,\ell}$ describes the ℓ th component of vector \mathbf{x}_k . The Euclidean norm of a vector $\mathbf{a}, \mathbf{a} \in \mathbb{C}^N$, is written as $\|\mathbf{a}\|$, and the absolute value of $b, b \in \mathbb{C}$, is $|b|$. $(\cdot)^H$ denotes the Hermitian transpose. The i th eigenvalue of a matrix \mathbf{Z} is denoted by $\mu_i(\mathbf{Z})$. The eigenvector which belongs to the i th eigenvalue of the matrix \mathbf{Z} is denoted by $\mathbf{v}_i(\mathbf{Z})$. We always assume that the eigenvalues are ordered in nondecreasing order such that $\mu_i(\mathbf{Z}) \leq \mu_{i+1}(\mathbf{Z})$. Moreover, the eigenvectors corresponding to the largest and smallest eigenvalues of a matrix \mathbf{Z} are specified as $\mathbf{v}_{\max}(\mathbf{Z})$ and $\mathbf{v}_{\min}(\mathbf{Z})$, respectively. The notation $\mathbf{Z} \succeq 0$ means that \mathbf{Z} is positive semidefinite. The rank and trace of a matrix \mathbf{Z} are given by $\text{rank}(\mathbf{Z})$ and $\text{tr}(\mathbf{Z})$, respectively. The orthogonal projector onto the column space of \mathbf{Z} is $\mathbf{\Pi}_Z := \mathbf{Z}(\mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{Z}^H$. The orthogonal projector onto the orthogonal complement of the column space of \mathbf{Z} is $\mathbf{\Pi}_Z^\perp := \mathbf{I} - \mathbf{\Pi}_Z$, where \mathbf{I} is an identity matrix. $\mathbb{E}(\cdot)$ denotes statistical expectation. The set of non-negative real numbers is \mathbb{R}_+ . The cardinality of a set \mathcal{K} is written as $|\mathcal{K}|$.

II. SYSTEM AND CHANNEL MODEL

We consider T transmitters and K receivers sharing the same spectral band. Define the set of transmitters as $\mathcal{T} := \{1, \dots, T\}$ and receivers as $\mathcal{K} := \{1, \dots, K\}$. Each transmitter sends useful information to at least one receiver. For transmitter $k, k \in \mathcal{T}$, let $\overline{\mathcal{K}}(k) \subseteq \mathcal{K}$ denote the set of its intended receivers for which useful information is sent to, and let $\underline{\mathcal{K}}(k) = \mathcal{K} \setminus \overline{\mathcal{K}}(k)$ be the set of its unintended receivers. Each transmitter k is equipped with N_k antennas, and each receiver with a single antenna. The quasi-static block flat-fading instantaneous channel vector from transmitter $k, k \in \mathcal{T}$, to receiver $\ell, \ell \in \mathcal{K}$, is denoted by $\mathbf{h}_{k\ell} \in \mathbb{C}^{N_k \times 1}$. The transmit covariance matrix of transmitter k is given as $\mathbf{Q}_k \in \mathbb{C}^{N_k \times N_k}$, $\mathbf{Q}_k \succeq 0$. We do not make any assumptions on the number of data streams applied at the transmitters. The basic model for the matched-filtered, symbol-sampled complex baseband data received at receiver ℓ is

$$y_\ell = \sum_{k=1}^T \mathbf{h}_{k\ell}^H \mathbf{Q}_k^{\frac{1}{2}} \mathbf{s}_k + n_\ell, \quad (1)$$

where \mathbf{s}_k is the symbols vector transmitted by transmitter k and n_ℓ are the noise terms which we model as independent and identically distributed (i.i.d.) complex Gaussian with zero mean and variance σ^2 .

Each transmitter has a total power constraint of $P := 1$ which leads to the constraint $\text{tr}(\mathbf{Q}_k) \leq 1$, $k \in \mathcal{T}$. Throughout, we define the signal to noise ratio (SNR) as $1/\sigma^2$. The feasible set of covariance matrices for transmitter k is defined as

$$\mathcal{S}_k := \{\mathbf{Q}_k \in \mathbb{C}^{N_k \times N_k} : \mathbf{Q}_k \succeq 0, \text{tr}(\mathbf{Q}_k) \leq 1\}. \quad (2)$$

Note that \mathcal{S}_k is compact and convex. We assume each transmitter has local CSI, i.e., it has perfect knowledge of the channel vectors only between itself and all receivers [7]. This ideal scenario serves as an upper bound to the more realistic case in which imperfect or partial CSI at the transmitters is available. Extensions in this direction are reported in [23], [24]. In these works, Pareto efficient transmission strategies are characterized for the two-user MISO IC with partial CSI at the transmitters.

A. Assumptions on Performance Measure

The performance measure of a system in an interference network is usually described by a utility function. The utility function associated with a receiver depends on the power gains originating from the transmitters in the network. Define the power gain achieved by transmitter k at a receiver ℓ as

$$x_{k,\ell}(\mathbf{Q}_k) = \mathbf{h}_{k\ell}^H \mathbf{Q}_k \mathbf{h}_{k\ell}, \quad (3)$$

where $x_{\ell}(\mathbf{Q}_k) \in \mathbb{R}_+$ since \mathbf{Q}_k is positive semidefinite. The utility function associated with a receiver ℓ is defined as $u_{\ell} : \mathbb{R}_+^T \rightarrow \mathbb{R}_+$, where T is the number of transmitters in the network.

Assumption 1: The utility function u_{ℓ} , $\ell \in \mathcal{K}$, has the following properties:

A. If $\ell \in \overline{\mathcal{K}}(k)$, then u_{ℓ} is monotonically increasing in the power gain from transmitter k , i.e.,

$$u_{\ell}(x_{1,\ell}(\mathbf{Q}_1), \dots, x_{T,\ell}(\mathbf{Q}_T)) \leq u_{\ell}(x_{1,\ell}(\mathbf{Q}_1), \dots, x_{k,\ell}(\widehat{\mathbf{Q}}_k), \dots, x_{T,\ell}(\mathbf{Q}_T)), \quad (4)$$

for $x_{k,\ell}(\mathbf{Q}_1) \leq x_{k,\ell}(\widehat{\mathbf{Q}}_k)$.

B. If $\ell \in \underline{\mathcal{K}}(k)$, then u_{ℓ} is monotonically decreasing in the power gain from transmitter k , i.e.,

$$u_{\ell}(x_{1,\ell}(\mathbf{Q}_1), \dots, x_{T,\ell}(\mathbf{Q}_T)) \geq u_{\ell}(x_{1,\ell}(\mathbf{Q}_1), \dots, x_{k,\ell}(\widehat{\mathbf{Q}}_k), \dots, x_{T,\ell}(\mathbf{Q}_T)), \quad (5)$$

for $x_{k,\ell}(\mathbf{Q}_k) \leq x_{k,\ell}(\widehat{\mathbf{Q}}_k)$. \square

Assumption 1 describes the settings where the performance measure at a receiver increases monotonically with increased power gain from intended transmitters and decreases monotonically with increased power gain from unintended transmitters. An example utility function which satisfies Assumption 1 is the signal to interference plus noise ratio (SINR).

The *utility region* is the set of all achievable utility tuples defined as:

$$\mathcal{U} := \{(u_1(x_{1,1}(\mathbf{Q}_1), \dots, x_{T,1}(\mathbf{Q}_T)), \dots, u_K(x_{1,K}(\mathbf{Q}_1), \dots, x_{T,K}(\mathbf{Q}_T))) : \mathbf{Q}_k \in \mathcal{S}_k, k \in \mathcal{T}\} \subset \mathbb{R}_+^K. \quad (6)$$

The efficient operating points in the utility region correspond to those in which it is impossible to improve the performance of one system without simultaneously degrading the performance of at least one other system. Such operating points are called Pareto optimal and are defined formally as follows.

Definition 1: A tuple $(u_1, \dots, u_K) \in \mathcal{U}$ is Pareto optimal if there is no other tuple $(u'_1, \dots, u'_K) \in \mathcal{U}$ such that $(u'_1, \dots, u'_K) \geq (u_1, \dots, u_K)$, where the inequality is component-wise and strict for at least one component. The set of all Pareto optimal operating points constitutes the *Pareto boundary* (\mathcal{PB}) of \mathcal{U} . \square

Next, we give an example setting where the systems' utility functions satisfy Assumption 1.

B. Example Setting

Consider two transmitters each using three transmit antennas, and three single antenna receivers as depicted in Fig. 1. The operation of the systems is as follows:

- **Broadcast Channel (BC):** Transmitter 1 transmits different useful data to receivers 1 and 2 simultaneously. We assume transmitter 1 chooses the transmit covariance matrices \mathbf{Q}_{11} with $\text{tr}(\mathbf{Q}_{11}) = p_{11}$ for receiver 1 and \mathbf{Q}_{12} with $\text{tr}(\mathbf{Q}_{12}) = p_{12}$ for receiver 2. Hence, transmitter 1 can be considered as two virtual transmitters², 11 and 12, coupled by the total power constraint, $p_{11} + p_{12} \leq 1$. The receivers are identified in the following receiver sets: $1 \in \overline{\mathcal{K}}(11), 1 \in \underline{\mathcal{K}}(12), 2 \in \overline{\mathcal{K}}(12), 2 \in \underline{\mathcal{K}}(11)$.
- **Multiple Access Channel (MAC):** Transmitters 12 and 2 send distinct useful information to receiver 2. Receiver 2 decodes the data from transmitter 12 and 2 successively. Thus, $2 \in \overline{\mathcal{K}}(12), 2 \in \overline{\mathcal{K}}(2)$.
- **Multicast:** Transmitter 2 sends common useful data in a multicast to receivers 2 and 3. The receivers are identified in the following receiver sets: $2 \in \overline{\mathcal{K}}(2), 3 \in \overline{\mathcal{K}}(2)$.
- **Interference Channel (IC):** Transmitter 2 induces interference on receiver 1, while transmitter 1 induces interference on receiver 3.

The receiver sets are summarized in Fig. 1, and the solid and dashed arrows refer to useful and not useful signal directions, respectively. The achievable rate at receiver 1 is

$$u_1(x_{11,1}(\mathbf{Q}_{11}), x_{12,1}(\mathbf{Q}_{12}), x_{2,1}(\mathbf{Q}_2)) = \log_2 \left(1 + \frac{\mathbf{h}_{11}^H \mathbf{Q}_{11} \mathbf{h}_{11}}{\sigma^2 + \mathbf{h}_{11}^H \mathbf{Q}_{12} \mathbf{h}_{11} + \mathbf{h}_{21}^H \mathbf{Q}_2 \mathbf{h}_{21}} \right), \quad (7)$$

²This transmission strategy is suboptimal, however less complex and more robust than dirty paper coding [25].

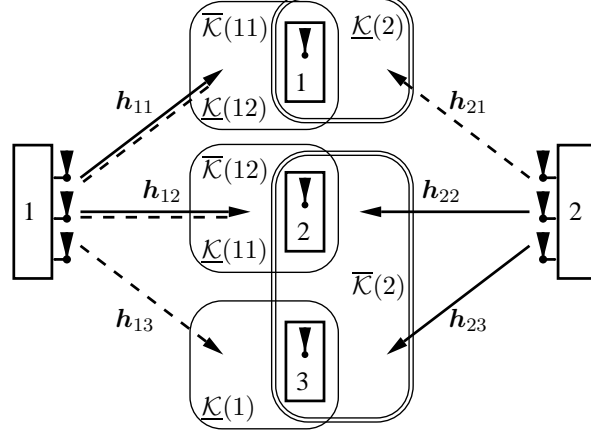


Fig. 1. An example setting for the described system model. There exist two transmitters, each equipped with three antennas, and three single antenna receivers. The solid arrows refer to the intended receivers of a transmitter, while the dashed arrows refer to interference directions.

which is monotonically increasing in $x_{11,1}(\mathbf{Q}_{11})$ and monotonically decreasing in the power gains from transmitters 12 and 2. The utility at receiver 2 is its sum capacity [26],

$$u_2(x_{11,2}(\mathbf{Q}_{11}), x_{12,2}(\mathbf{Q}_{12}), x_{2,2}(\mathbf{Q}_2)) = \log_2 \left(1 + \frac{\mathbf{h}_{12}^H \mathbf{Q}_{12} \mathbf{h}_{12} + \mathbf{h}_{22}^H \mathbf{Q}_2 \mathbf{h}_{22}}{\sigma^2 + \mathbf{h}_{12}^H \mathbf{Q}_{11} \mathbf{h}_{12}} \right), \quad (8)$$

which is monotonically increasing in $x_{12,2}(\mathbf{Q}_{12})$ and $x_{2,2}(\mathbf{Q}_2)$. The utility function at receiver 3 is the achievable rate³,

$$u_3(x_{11,3}(\mathbf{Q}_{11}), x_{12,3}(\mathbf{Q}_{12}), x_{2,3}(\mathbf{Q}_2)) = \log_2 \left(1 + \frac{\mathbf{h}_{23}^H \mathbf{Q}_2 \mathbf{h}_{23}}{\sigma^2 + \mathbf{h}_{13}^H \mathbf{Q}_{11} \mathbf{h}_{13} + \mathbf{h}_{13}^H \mathbf{Q}_{12} \mathbf{h}_{13}} \right). \quad (9)$$

which is monotonically increasing in $x_{2,3}(\mathbf{Q}_2)$. The utility functions in (7)-(9) satisfy properties A and B in Assumption 1. We return to this setting in Section IV-A after we formalize the framework for efficient beamforming and resource allocation.

III. POWER GAIN REGION

In this section, a single transmitter $k, k \in \mathcal{T}$, is considered along with all K receivers. The subscript k in all terms referring to the single transmitter is omitted for convenience. For example, \mathbf{h}_ℓ is written instead of $\mathbf{h}_{k\ell}$. The receiver sets of the transmitter are written as $\bar{\mathcal{K}}$ and $\underline{\mathcal{K}}$ instead of $\bar{\mathcal{K}}(k)$ and $\underline{\mathcal{K}}(k)$.

³Note that the transmission rate at transmitter 2 has to be chosen such that both receiver 2 and 3 can decode the data successfully. We do not consider this requirement in (8) and (9) since this is beyond the scope of this paper. However, these rates can be achieved using rateless coding [27], [28].

We return to include the indication to the transmitter in the next sections when multiple transmitters are considered. Here, we study the transmission effects of a single transmitter on all existing receivers. Thereby, we characterize its transmit covariance matrices that are relevant for its efficient operation in the multiuser system. For all feasible transmit covariance matrices, a power gain-region of a single transmitter is the set that consists of all joint power gains achievable at the receivers. The power gain-region of a transmitter is defined as

$$\Omega := \{(x_1(\mathbf{Q}), \dots, x_K(\mathbf{Q})) : \mathbf{Q} \in \mathcal{S}\} \subset \mathbb{R}_+^K, \quad (10)$$

where \mathcal{S} is defined in (2). Note that $x_1(\mathbf{Q}), \dots, x_K(\mathbf{Q})$ are the main elements of the utility functions in (7)-(9). An important property of the gain-region Ω is its convexity. Having this property is convenient for characterizing the points on its boundary using simple programming problems based on the Hyperplane Separation theorem [29, Theorem 1.3].

Lemma 1: The set Ω in (10) is a compact and convex set.

Proof: The proof is provided in Appendix A. ■

Points that lie on the boundary of the gain-region Ω are of interest since these points characterize extreme power gains achievable at the receivers. At these points, the transmitter cannot increase the power gain in one direction of the gain-region without decreasing the power gain in any other direction. We give an example later in this section which further clarifies the importance of the boundary points of the gain-region for efficient operation of a transmitter. Next, we formalize the boundary of the set Ω following the definitions in [5]. There, these definitions were used to derive the solution of monotonic optimization problems [30], [29].

Definition 2: A point $\mathbf{y} \in \mathbb{R}_+^n$ is called upper boundary point of a compact convex set \mathcal{C} if $\mathbf{y} \in \mathcal{C}$ while the set $\{\mathbf{y}' \in \mathbb{R}_+^n : \mathbf{y}' > \mathbf{y}\} \subset \mathbb{R}_+^n \setminus \mathcal{C}$, where the inequality in $\mathbf{y}' > \mathbf{y}$ is componentwise. The set of upper boundary points of \mathcal{C} is denoted by $\partial^+ \mathcal{C}$. □

Definition 2 describes only one boundary part of a compact convex set. The straightforward extension to describe all boundary parts of this set is to define its upper boundary in direction $\mathbf{e}, \mathbf{e} \in \{-1, +1\}^n$. For this purpose, we first need the following definition.

Definition 3: A vector \mathbf{x} dominates a vector \mathbf{y} in direction \mathbf{e} , written as $\mathbf{x} \geq^{\mathbf{e}} \mathbf{y}$, if $x_\ell e_\ell \geq y_\ell e_\ell$ for all $\ell, 1 \leq \ell \leq n$, and the inequality has at least one strict inequality. □

Definition 4: A point $\mathbf{y} \in \mathbb{R}_+^n$ is called upper boundary point of a compact convex set \mathcal{C} in direction \mathbf{e} if $\mathbf{y} \in \mathcal{C}$ while the set $\{\mathbf{y}' \in \mathbb{R}_+^n : \mathbf{y}' \geq^{\mathbf{e}} \mathbf{y}\} \subset \mathbb{R}_+^n \setminus \mathcal{C}$. We denote the set of upper boundary points in direction \mathbf{e} as $\partial^{\mathbf{e}} \mathcal{C}$. □

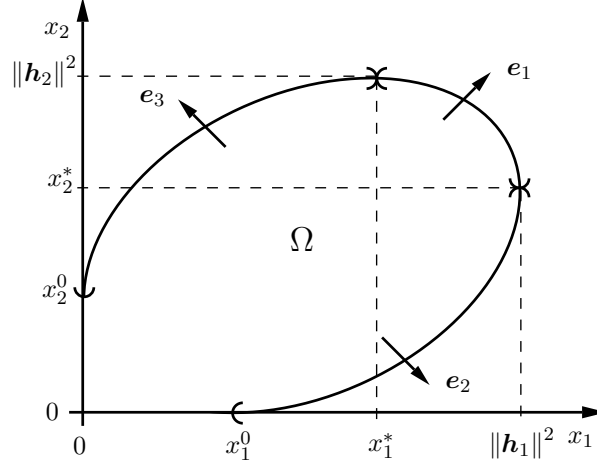


Fig. 2. An illustration of a two-dimensional gain-region and its upper boundaries in directions $e_1 = [1, 1]$, $e_2 = [1, -1]$, and $e_3 = [-1, 1]$.

An illustration of a two-dimensional power gain-region is given in Fig. 2. The direction vectors e_1, e_2 , and e_3 refer to three different parts of the boundary. For the choice $e = 1$, the upper boundary in direction e is the usual upper boundary as in Definition 2, i.e., $\partial^+ \Omega = \partial^1 \Omega$.

We draw an example to illustrate the importance of the boundary parts of the gain-region. Assume as in Fig. 2, there exist two single antenna receivers and a transmitter. Assume, receiver 1 is the intended receiver of the transmitter, i.e. $1 = \overline{\mathcal{K}}$, and receiver 2 is its unintended receiver such that $2 = \underline{\mathcal{K}}$. For efficient operation in the setting, the transmitter is interested in maximizing its power gain at receiver 1 and also interested in minimizing the power gain on receiver 2. If we seek to characterize the set of efficient transmission strategies of the transmitter, then the boundary part corresponding to $e_2 = [1, -1]$ is relevant. Any power gain tuple which is inside the power gain-region and does not lie on its boundary is not relevant, since the transmitter can increase the power gain to its intended receiver without changing the power gain to the unintended receiver. The direction vector e_2 corresponds to the boundary part which includes the maximum achievable power gain at receiver 1 and also the minimum achievable power gain at receiver 2. These extreme points correspond to MRT and ZF transmission strategies, respectively. According to Fig. 2, MRT achieves $\|h_1\|^2$ power gain at the first receiver and x_2^* at the second receiver. On the other hand, ZF transmission achieves zero power gain at the second receiver and x_1^0 at the first receiver. From this example, it can be observed that the boundary part which corresponds to the transmitters efficient strategy set is the one where the direction vector has positive component corresponding to the intended receiver and negative component corresponding to the unintended receiver.

Since a single transmitter has at least one intended receiver, otherwise it will not operate, the direction vector with all components equal to -1 is not of interest. Therefore, we define the set of feasible directions

$$\mathcal{E} := \{-1, 1\}^K \setminus \{-1\}^K. \quad (11)$$

Next, we study the transmission strategies that achieve the boundary points in $\partial^e \Omega$, $e \in \mathcal{E}$.

The relation between the number of existing receivers and the number of available antennas at the transmitter is important to distinguish whether power control is needed to achieve all boundary parts of the power gain-region. First, we consider the case where the number of transmit antennas is greater than or equal to the number of existing receivers K . For this case, a comprehensive study of the gain-region is given due to which we gain a link and some insights to a mathematical field of research in matrix analysis. Afterwards, we consider the case in which the number of antennas is strictly less than K . This case is addressed more briefly since the tools needed for the analysis are similar to those in the first case.

A. The number of transmit antennas satisfies $N \geq K$

In this section, we assume $N \geq K$ and the channel vectors to the receivers are independently distributed. These assumptions lead to linear independence of the channel vectors with probability one. In this case, it is possible to achieve power gain on one receiver and simultaneously null the power gain at the remaining receivers. Therefore, the gain-region has boundary points that lie on each axis as the points x_1^0 and x_2^0 in Fig. 2. As a result, all upper boundary points of the gain-region are achieved with full power. In case the number of antennas at the transmitter is strictly larger than K , the gain-region has part of its boundary the points between x_1^0 and the origin.

Lemma 2: Transmit covariance matrices from the set

$$\hat{\mathcal{S}} := \{\mathbf{Q} : \mathbf{Q} \succeq 0, \text{tr}(\mathbf{Q}) = 1\}, \quad (12)$$

achieve all points in $\partial^e \Omega$, $e \in \mathcal{E}$.

Proof: The proof is provided in Appendix B. ■

Lemma 2 supports the result in [31], where it is shown that all Pareto efficient operating points in the MISO IC correspond to full power transmission when the number of antennas is larger than or equal to the number of receivers. The power gain-region achieved with full power transmit covariance matrices is defined as

$$\hat{\Omega} := \{(x_1(\mathbf{Q}), \dots, x_K(\mathbf{Q})) : \mathbf{Q} \in \hat{\mathcal{S}}\} \subset \mathbb{R}_+^K. \quad (13)$$

Since the power gain at a receiver ℓ can equivalently be formulated as $x_\ell(\mathbf{Q}) = \mathbf{h}_\ell^H \mathbf{Q} \mathbf{h}_\ell = \text{tr}(\mathbf{Q} \mathbf{h}_\ell \mathbf{h}_\ell^H)$, the set $\widehat{\Omega}$ in (13) is rewritten as

$$\widehat{\Omega} = \{(\text{tr}(\mathbf{Q} \mathbf{h}_1 \mathbf{h}_1^H), \dots, \text{tr}(\mathbf{Q} \mathbf{h}_K \mathbf{h}_K^H)) : \mathbf{Q} \succeq 0, \text{tr}(\mathbf{Q}) = 1\}. \quad (14)$$

This set is referred to in [32] as the *joint field of values* of the set of matrices $\mathbf{h}_1 \mathbf{h}_1^H, \dots, \mathbf{h}_K \mathbf{h}_K^H$. The set $\widehat{\Omega}$ is compact and convex⁴. The next result shows that the boundary of $\widehat{\Omega}$ in any direction $\mathbf{e} \in \mathcal{E}$ can be achieved with rank-1 transmit covariance matrices.

Lemma 3: Transmit covariance matrices from the set

$$\widetilde{\mathcal{S}} := \{\mathbf{Q} : \mathbf{Q} \in \widehat{\mathcal{S}}, \text{rank}(\mathbf{Q}) = 1\}, \quad (15)$$

achieve all points in $\partial^e \widehat{\Omega}$, $\mathbf{e} \in \mathcal{E}$.

Proof: The proof is provided in Appendix C. ■

Lemma 3 supports the result in [35] as a special case, where it is shown that single-stream beamforming is optimal in the MISO IC to achieve Pareto efficient operating points. Accordingly, efficient transmission strategies can be described by beamforming vectors. The next theorem characterizes the beamforming vectors that achieve the upper boundary of the set $\widehat{\Omega}$ in a specific direction \mathbf{e} .

Theorem 1: All upper boundary points of the set $\widehat{\Omega}$ in direction $\mathbf{e} \in \mathcal{E}$ can be achieved by

$$\mathbf{w}(\boldsymbol{\lambda}) = \mathbf{v}_{\max} \left(\underbrace{\sum_{\ell=1}^K \lambda_\ell e_\ell \mathbf{h}_\ell \mathbf{h}_\ell^H}_{\mathbf{Z}} \right), \quad (16)$$

with

$$\boldsymbol{\lambda} \in \boldsymbol{\Lambda} := \left\{ \boldsymbol{\lambda} \in [0, 1]^K : \sum_{\ell=1}^K \lambda_\ell = 1 \right\}. \quad (17)$$

Proof: The proof is provided in Appendix D. ■

The interesting observation from Theorem 1 is that all upper boundary points in direction \mathbf{e} of the K -dimensional gain-region can be achieved by a parametrization using $K - 1$ real parameters. Note that the components of the direction vector \mathbf{e} in Theorem 1 can have positive and negative components. For the particular choices $\lambda_\ell = 0$ for all ℓ such that $e_\ell = +1$, the largest eigenvalue of \mathbf{Z} in (16) is zero. In this case the largest eigenvalue of \mathbf{Z} can also have geometric multiplicity larger than one, i.e., there

⁴The set $\widehat{\Omega}$ is the convex hull of the *joint numerical range* of the matrices $\mathbf{h}_1 \mathbf{h}_1^H, \dots, \mathbf{h}_K \mathbf{h}_K^H$. The joint numerical range of a set of K matrices $\mathbf{A}_1, \dots, \mathbf{A}_K$ is defined as [33]: $\mathcal{W}(\mathbf{A}_1, \dots, \mathbf{A}_K) := \{(\mathbf{z}^H \mathbf{A}_1 \mathbf{z}, \dots, \mathbf{z}^H \mathbf{A}_K \mathbf{z}) : \mathbf{z}^H \mathbf{z} = 1\}$. In our case, $\mathbf{A}_1 = \mathbf{h}_1 \mathbf{h}_1^H, \dots, \mathbf{A}_K = \mathbf{h}_K \mathbf{h}_K^H$. Convexity of the set $\mathcal{W}(\mathbf{A}_1, \dots, \mathbf{A}_K)$ is not always satisfied. Conditions for its convexity are studied in [34].

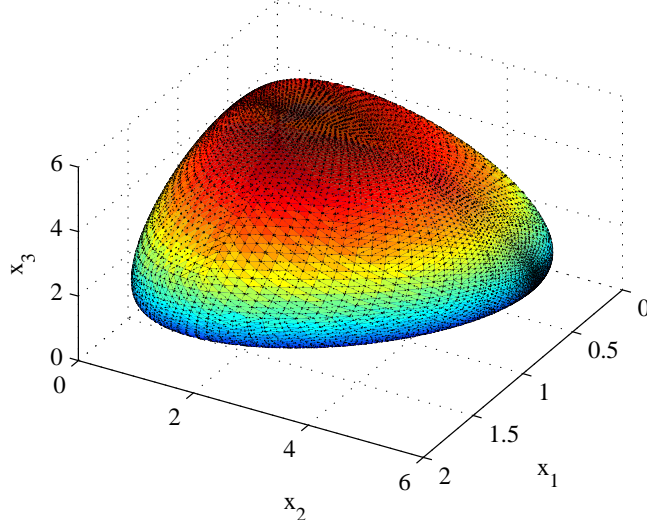


Fig. 3. An illustration of a three dimensional gain-region where the transmitter uses three antennas.

exist multiple linearly independent eigenvectors associated with that eigenvalue. For example, assume $N = 3, K = 2, e_1 = +1$, and $e_2 = -1$. For $\lambda_1 = 0$ and $\lambda_2 = 1$, the matrix \mathbf{Z} in Theorem 1 is equal to $-\mathbf{h}_2\mathbf{h}_2^H$ which is rank-1 with the largest eigenvalue having geometric multiplicity two. In this case, we choose an eigenvector which lies in the span of $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K]$.

In Fig. 3, a three dimensional gain-region is plotted where the transmitter uses three transmit antennas. The boundary points are calculated by generating the beamforming vectors characterized in Theorem 1. For each boundary part, two real-valued parameters are required, and these are varied between zero and one with a step-length of 0.02. The gain-region is shown to have a convex boundary. Moreover, since it is possible to null out the power gain at two receivers simultaneously, the boundary of the gain-region touches each axis in one point.

B. The number of transmit antennas satisfies $N < K$

In case the number of antennas at the transmitter satisfies $N < K$, it is not possible for the transmitter to choose a full power transmission strategy which nulls out the power gain at all except one receiver. Hence, the transmitter's ZF strategy would be to allocate no transmit power. This reveals that in order to achieve all boundary points of the power gain-region the transmitter has to reduce its transmission power. This is the reason why we study this case separately.

We start by assuming that the transmit covariance matrices are chosen from the set \mathcal{S} in (2). The corresponding power gain-region is compact and convex according to Lemma 1. Hence the following

programming problem, similar to the one formulated in (35) in Appendix C, achieves the boundary points of Ω in direction $\mathbf{e} \in \mathcal{E}$:

$$\begin{aligned} & \text{maximize} \quad \sum_{\ell=1}^K \lambda_{\ell} e_{\ell} \mathbf{h}_{\ell}^H \mathbf{Q} \mathbf{h}_{\ell} \\ & \text{subject to} \quad \text{tr}(\mathbf{Q}) \leq 1, \quad \mathbf{Q} \succeq 0, \end{aligned} \quad (18)$$

where $\boldsymbol{\lambda} \in \boldsymbol{\Lambda}$ is defined in (17). As in Appendix C, the problem in (18) can be equivalently written as

$$\begin{aligned} & \text{maximize} \quad \sum_{\ell=1}^N \mu_{\ell}(\mathbf{Q}) \underbrace{\mu_{\ell} \left(\sum_{\ell=1}^K \lambda_{\ell} e_{\ell} \mathbf{h}_{\ell} \mathbf{h}_{\ell}^H \right)}_{\mathbf{Z}} \\ & \text{subject to} \quad \sum_{\ell=1}^N \mu_{\ell}(\mathbf{Q}) \leq 1, \quad \mu_{\ell}(\mathbf{Q}) \geq 0, \quad \text{for all } \ell = 1, \dots, N \end{aligned} \quad (19)$$

The solution of this problem is $\mu_N(\mathbf{Q}) = p, p \in [0, 1]$, and $\mu_{\ell}(\mathbf{Q}) = 0$ for $\ell \neq N$. Hence, the optimal transmit covariance matrices \mathbf{Q} are rank-1. In addition, the optimal power allocation is to allocate power only in the direction of the eigenvector corresponding to the largest eigenvalue of \mathbf{Z} . Hence the formulation in Theorem 1 is still valid to determine the beamforming vectors that achieve the boundary points of Ω . However, power control is to be applied on specific beamforming vectors. In order to maximize the problem in (19), we choose $p = 1$ if the largest eigenvalue of \mathbf{Z} is strictly larger than zero, and $p = 0$ if the largest eigenvalue of \mathbf{Z} is strictly less than zero. It can be easily checked that \mathbf{Z} can be negative definite according to Weyl's eigenvalue inequality in [36, Theorem 4.3.7]. We give a simple example to illustrate this case. Assume $N = 2, K = 3, e_1 = +1$, and $e_2 = e_3 = -1$. For $\lambda_1 = 0$ and $\lambda_2, \lambda_3 > 0$, then $\mathbf{Z} = -\lambda_2 \mathbf{h}_2 \mathbf{h}_2^H - \lambda_3 \mathbf{h}_3 \mathbf{h}_3^H$ which is negative definite. In case the largest eigenvalue of \mathbf{Z} is zero, all feasible power allocations should be adopted in order to achieve the boundary of the gain region. For the same example given above and for $\mu(\mathbf{Z}) = 0$, the weighted sum gains in (39) from Appendix D formulates to $\lambda_1 x_1(\mathbf{w}_1 \mathbf{w}_1^H) = \lambda_2 x_2(\mathbf{w}_1 \mathbf{w}_1^H) + \lambda_3 x_3(\mathbf{w}_1 \mathbf{w}_1^H)$. This equation, with the variables in $\boldsymbol{\lambda}$, describes the plane that touches the boundary of the power gain-region at $\mathbf{x}(\mathbf{w}_1 \mathbf{w}_1^H)$ and passes through the origin. Choosing a power allocation $p \in [0, 1]$ achieves points on the segment connecting the origin and $\mathbf{x}(\mathbf{w}_1 \mathbf{w}_1^H)$ which correspond to boundary points on the power gain-region. The power allocations that achieve the boundary points of the power gain-region Ω in direction $\mathbf{e} \in \mathcal{E}$ are summarized as follows:

$$p = 1 \text{ for } \mu_N(\mathbf{Z}) > 0, \quad p \in [0, 1] \text{ for } \mu_N(\mathbf{Z}) = 0, \quad p = 0 \text{ for } \mu_N(\mathbf{Z}) < 0. \quad (20)$$

where \mathbf{Z} is given in (19).

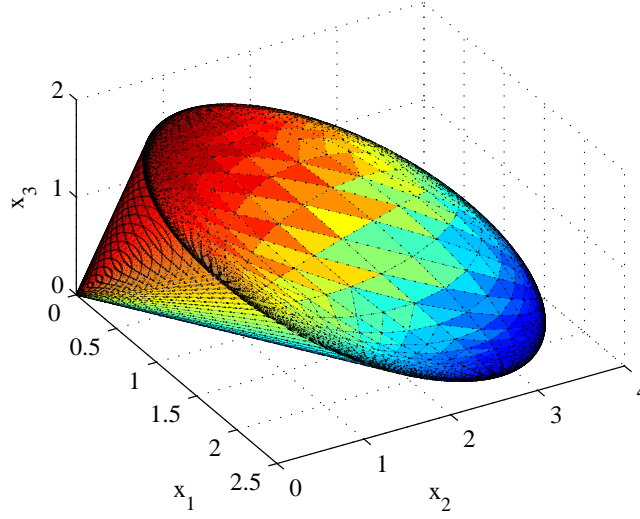


Fig. 4. An illustration of a three dimensional gain-region where the transmitter uses two antennas.

In Fig. 4, a three dimensional gain-region is plotted where two antennas are utilized at the transmitter. In comparison to the plot in Fig. 3, the region looks like a cone whose vertex is at the origin. The outermost boundary which is not flat is attained with full power transmission. The flat surfaces correspond to the case for which the transmission power is varied between zero and one.

IV. PARETO BOUNDARY CHARACTERIZATION

In this section, all T transmitters are considered again. The analysis of the single transmitter case in the previous section builds the framework to determine the beamforming vectors for each transmitter that are relevant for Pareto optimal operation. Each transmitter $k, k \in \mathcal{T}$, is associated with a power gain-region Ω_k . The sets of intended and unintended receivers of a transmitter k are $\overline{\mathcal{K}}(k)$ and $\underline{\mathcal{K}}(k)$, respectively. In a general network, efficient operation of the systems requires the transmitters to maximize the power gain at intended receivers and simultaneously minimize the power gain at unintended receivers. In order to achieve Pareto optimal operating points a combination of these objectives is required at each transmitter. The gain-region characterization in the previous section, illustrates the effects of the beamforming vectors on the power gains achieved jointly at the receivers. Next, we formalize the transmit beamforming vectors of each transmitter that are relevant to achieve Pareto optimal points in the utility region.

Theorem 2: All Pareto optimal points in the utility region \mathcal{U} can be achieved by beamforming vectors

$$\mathbf{w}_k(\boldsymbol{\lambda}_k) = p_k \mathbf{v}_{max} \left(\underbrace{\sum_{\ell=1}^K \lambda_{k,\ell} e_{k,\ell} \mathbf{h}_{k,\ell} \mathbf{h}_{k,\ell}^H}_{\mathbf{Z}_k} \right), \quad (21)$$

with $\boldsymbol{\lambda}_k \in \boldsymbol{\Lambda}$ defined in (17) and

$$e_{k,\ell} = \begin{cases} +1 & \ell \in \overline{\mathcal{K}}(k) \\ -1 & \ell \in \underline{\mathcal{K}}(k) \end{cases}, \quad (22)$$

$$p_k = \begin{cases} 1 & \mu_N(\mathbf{Z}_k) > 0 \\ [0, 1] & \mu_N(\mathbf{Z}_k) = 0 \\ 0 & \mu_N(\mathbf{Z}_k) < 0 \end{cases}. \quad (23)$$

Proof: The proof is provided in Appendix E. ■

Theorem 2 characterizes the transmission strategies that achieve Pareto optimal points. The number of real-valued parameters, $\boldsymbol{\lambda}$ and p , that are required to characterize these beamforming vectors is: (i) $T(K-1)$ in case no power control is needed. This correspond to the case when $N_k > |\underline{\mathcal{K}}(k)|$. (ii) TK in case power control is needed corresponding to the case $N_k \leq |\underline{\mathcal{K}}(k)|$. All these parameters take values between zero and one. The direction vector in (22) can be determined since each transmitter k knows its intended and unintended receiver sets, $\overline{\mathcal{K}}(k)$ and $\underline{\mathcal{K}}(k)$ respectively. This direction vector specifies the relevant boundary of the transmitter's power gain-region.

A. Example Revisited

The example in Section II-B will be continued to find the beamforming vectors that achieve Pareto optimal points in the corresponding utility region. The choice of the beamforming vectors with unit norm for transmitters 11, 12 and 2 which achieve Pareto optimal points are characterized in Theorem 2. Since the number of antennas at each transmitter is strictly larger than the number of unintended receivers, all transmitters should transmit at full power according to (23). The power allocation p_{11} and p_{12} on the beamforming vectors \mathbf{w}_{11} and \mathbf{w}_{12} , respectively, is varied such $p_{11} + p_{12} = 1$. The power allocation parameters p_{11} and p_{12} can be expressed as q and $1-q$, respectively, with $q \in [0, 1]$. The characterization in Theorem 2 leads to the following nonnegative real-valued parameters:

- For transmitter 11: $\lambda_{11,1}, \lambda_{11,2}, \lambda_{11,3}$, with $\lambda_{11,1} + \lambda_{11,2} + \lambda_{11,3} = 1$.
- For transmitter 12: $\lambda_{12,1}, \lambda_{12,2}, \lambda_{12,3}$, with $\lambda_{12,1} + \lambda_{12,2} + \lambda_{12,3} = 1$.

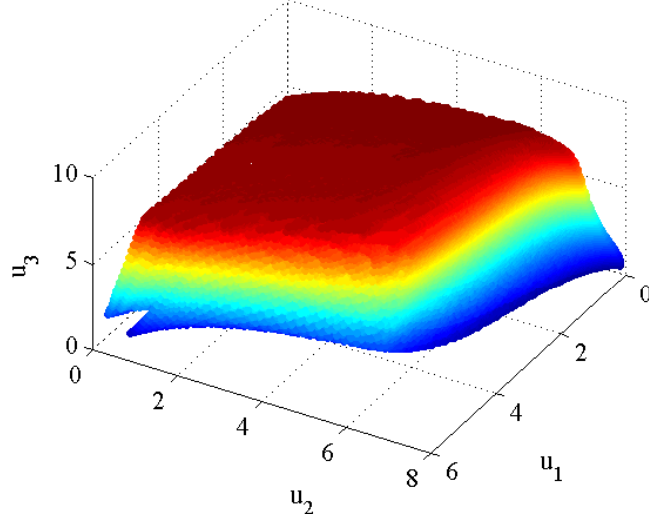


Fig. 5. Pareto boundary of the utility region of the setting described in Section IV-A with SNR = 15 dB and $N = 3$.

- For transmitters 11 and 12: $q \in [0, 1]$.
- For transmitter 2: $\lambda_{2,1}, \lambda_{2,2}, \lambda_{2,3}$, with $\lambda_{2,1} + \lambda_{2,2} + \lambda_{2,3} = 1$.

All seven required parameters are in the interval $[0, 1]$, and the plot in Fig. 5 is obtained by varying these in a grid with 0.05 step-length. The points obtained include the points that lie on the Pareto boundary of the utility region which satisfy Definition 1. Points corresponding to weak Pareto optimality are not achieved by the parametrization in Theorem 2 and thus not included in the plot. Weak Pareto optimality is defined as in Definition 1 except that the corresponding inequality is not strict. Weak Pareto optimal points complete the shape of the utility region by orthogonally projecting each Pareto optimal point in Fig. 5 onto the coordinate surfaces.

In the next section, we discuss two special applications of the developed framework.

V. APPLICATIONS

A. Multiple-input Single-output Interference Channel

The K -users MISO IC consists of K transmitter-receiver pairs, where each receiver has an intended transmitter while all other transmitters induce interference on this receiver. We consider single-user decoding, i.e., interference is treated as additive noise at each receiver. For a given set of beamforming vectors $\{\mathbf{w}_1, \dots, \mathbf{w}_K\}$, the following rate is achievable at receiver k , by using codebooks approaching

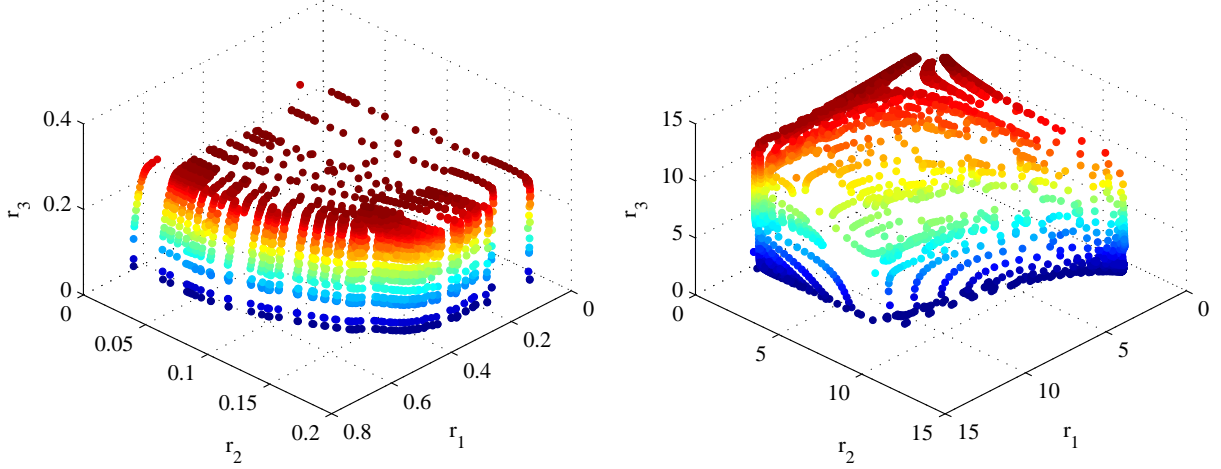


Fig. 6. Pareto boundary of a three user MISO IC rate-region, where $N = 3$. In the figure on the left, SNR = -10 dB, and in the figure on the right, SNR = 30 dB.

Gaussian ones:

$$u_k(\mathbf{w}_1 \mathbf{w}_1^H, \dots, \mathbf{w}_K \mathbf{w}_K^H) = \log_2 \left(1 + \frac{|\mathbf{w}_k^H \mathbf{h}_{kk}|^2}{\sum_{\ell \neq k} |\mathbf{w}_\ell^H \mathbf{h}_{\ell k}|^2 + \sigma^2} \right). \quad (24)$$

This utility function satisfies properties A and B in Assumption 1 which leads to the following receiver sets for each transmitter k : $\overline{\mathcal{K}}(k) = \{k\}$, and $\underline{\mathcal{K}}(k) = \mathcal{K} \setminus \{k\}$. All points on the Pareto boundary of the achievable rate-region of the MISO IC can be reached by beamforming vectors as given in Theorem 2. In [13], a characterization of the beamforming vectors that achieve the Pareto boundary of the achievable rate-region is provided by a complex linear combination of the MRT and ZF strategies. The parametrization in Theorem 2 is real-valued with the same number of parameters, thus of lower dimension.

The Pareto boundary of a three user MISO IC rate-region is plotted in Fig. 6 (left) for SNR = -10 dB and in Fig. 6 (right) for SNR = 30 dB. The generated points correspond to the beamforming vectors characterized in Theorem 2. The real-valued parameters are varied in a 0.05 step-length. Since we are only interested in revealing the Pareto boundary of the achievable rate region, we do the following. We randomly choose ten thousand generated points and remove all points that are dominated by these. In other words, for a randomly chosen rate tuple, all points corresponding to joint rates less than the chosen one are removed. In addition, the algorithm provided in [37] is applied. The algorithm reduces the number of plotted points in removing the points that are not visible from the viewed angle of the figure. This algorithm further reduces the complexity of rendering the generated points.

Next, we identify special operation points of each transmitter.

1) *Maximum ratio transmission*: Maximum ratio transmission for transmitter k corresponds obviously to the parameters $\lambda_{k,\ell} = 1$ for $k = \ell$ and $\lambda_{k,\ell} = 0$ for $k \neq \ell$. This transmission strategy is the unique Nash equilibrium (NE) strategy of each transmitter k , which is the outcome of a noncooperative game between the users.

2) *Zero-forcing transmission*: Zero-forcing transmission is characterized with the following lemma:

Lemma 4: Zero Forcing transmission for transmitter k corresponds to the parameters $\lambda_{k,\ell} = 0$ for $k = \ell$ and $\lambda_{k,\ell} > 0$ for $k \neq \ell$.

Proof: The proof is provided in Appendix F. ■

The two-user MISO IC special case has an appealing form in terms of the parametrization of the Pareto boundary of the achievable rate-region. The efficient beamforming vectors are a linear combination of the MRT and ZF strategies [13, Corollary 2]. These two strategies have the interpretation of a transmitter being either selfish or altruistic [4], [31]. The efficient beamforming vectors of transmitter $k \in \{1, 2\}$ are given as [13, Corollary 2]

$$\mathbf{w}_k(\hat{\lambda}_k) = \frac{\hat{\lambda}_k \mathbf{w}_k^{\text{MRT}} + (1 - \hat{\lambda}_k) \mathbf{w}_k^{\text{ZF}}}{\left\| \hat{\lambda}_k \mathbf{w}_k^{\text{MRT}} + (1 - \hat{\lambda}_k) \mathbf{w}_k^{\text{ZF}} \right\|}, \quad (25)$$

where $\hat{\lambda}_k \in [0, 1]$, $\mathbf{w}_k^{\text{MRT}} = \frac{\mathbf{h}_{kk}}{\|\mathbf{h}_{kk}\|}$, and $\mathbf{w}_k^{\text{ZF}} = \frac{\Pi_{\mathbf{h}_{k\ell}}^\perp \mathbf{h}_{kk}}{\|\Pi_{\mathbf{h}_{k\ell}}^\perp \mathbf{h}_{kk}\|}$, $k \neq \ell$. Next, we prove that the parametrization in (25) has the same set of strategies as in Theorem 2 for $K = 2$. For this case, the eigenvalue equation for the hermitian matrix in Theorem 1 is written as $(\lambda_1 \mathbf{h}_{11} \mathbf{h}_{11}^H - (1 - \lambda_1) \mathbf{h}_{12} \mathbf{h}_{12}^H) \mathbf{w}_1 = \mu \mathbf{w}_1$. This equation can be equivalently formulated to $\lambda_1 \|\mathbf{h}_{11}\|^2 \frac{\mathbf{h}_{11} \mathbf{h}_{11}^H}{\|\mathbf{h}_{11}\|^2} \mathbf{w}_1 - (1 - \lambda_1) \|\mathbf{h}_{12}\|^2 \frac{\mathbf{h}_{12} \mathbf{h}_{12}^H}{\|\mathbf{h}_{12}\|^2} \mathbf{w}_1 = \mu \mathbf{w}_1$. Adding $(1 - \lambda_1) \|\mathbf{h}_{12}\|^2 \mathbf{w}_1$ on both sides of the equation leads to

$$\left(\lambda_1 \|\mathbf{h}_{11}\|^2 \Pi_{\mathbf{h}_{11}} + (1 - \lambda_1) \|\mathbf{h}_{12}\|^2 \Pi_{\mathbf{h}_{12}}^\perp \right) \mathbf{w}_1 = (\mu + (1 - \lambda_1) \|\mathbf{h}_{12}\|^2) \mathbf{w}_1. \quad (26)$$

The LHS of the equation states that the principal eigenvector is a linear combination of its orthogonal projection on \mathbf{h}_{11} and the orthogonal projection onto the orthogonal complement of \mathbf{h}_{12} . Since the largest eigenvalue μ is larger or equal to zero, the weight in the RHS of (26) is always positive. Hence, the optimal set of beamforming vectors can be equivalently characterized by (25).

B. Noncooperative Underlay Cognitive Radio

In an underlay cognitive radio scenario, secondary users can share the communication resources with primary users under the condition of not imposing quality of service (QoS) degradation to the primary systems. A limited QoS degradation to the primary users is described by interference temperature

constraints (ITC) [38]. When no interference on the primary users is allowed, the constraint is said to be a null-shaping constraint. Motivated by the concept of null-shaping constraints [39], the next result gives an alternative characterization of efficient transmission strategies. In order to be able to fulfill the null-shaping constraints, the number of applied antennas at the transmitter has to be greater than or equal to the number of primary receivers. We assume there exists virtual single-antenna primary receivers, and consider the efficient design of the null-shaping constraints. The following result has been presented in [40] without proof.

Corollary 1: Assume $N_k \geq K$ and define the matrix

$$Z_k(\boldsymbol{\lambda}_k) = \left[z_1(\boldsymbol{\lambda}_k), \dots, z_{|\underline{K}(k)|}(\boldsymbol{\lambda}_k), z_{N_k - |\overline{K}(k)| + 1}(\boldsymbol{\lambda}_k), \dots, z_{N_k}(\boldsymbol{\lambda}_k) \right], \quad (27)$$

where

$$z_i(\boldsymbol{\lambda}_k) = \mathbf{v}_i \left(\sum_{\ell=1}^K \lambda_{k,\ell} e_{k,\ell} \mathbf{h}_{k\ell} \mathbf{h}_{k\ell}^H \right), \quad (28)$$

with $\boldsymbol{\lambda}_k \in \boldsymbol{\Lambda}$ defined in (17) and $e_{k,\ell}$ defined in (22). All points on the Pareto boundary of the utility region \mathcal{U} can be reached by beamforming vectors

$$\mathbf{w}_k(\boldsymbol{\lambda}_k) = \frac{\boldsymbol{\Pi}_{Z_k(\boldsymbol{\lambda}_k)}^\perp \mathbf{h}_{k\ell}}{\|\boldsymbol{\Pi}_{Z_k(\boldsymbol{\lambda}_k)}^\perp \mathbf{h}_{k\ell}\|}, \quad \ell \in \overline{K}(k). \quad (29)$$

Proof: The proof is provided in Appendix G. ■

In Corollary 1, the design of the null-shaping constraints is given in (27), and the efficient transmission strategies are given in (29). Here, $K - 1$ null-shaping constraints are to be applied on each transmitter, and the number of required real-valued parameters is the same as in Theorem 2. Hence, the complexity of parameterizing the efficient beamforming vectors is the similar in Corollary 1 and Theorem 2.

The form of the transmission strategy in (29) has a relevant interpretation. Consider a MISO IC setting as in the previous section. Rewriting (29) for transmitter k gives $\mathbf{w}_k(\boldsymbol{\lambda}_k) = \frac{\boldsymbol{\Pi}_{Z_k(\boldsymbol{\lambda}_k)}^\perp \mathbf{h}_{kk}}{\|\boldsymbol{\Pi}_{Z_k(\boldsymbol{\lambda}_k)}^\perp \mathbf{h}_{kk}\|}$. This transmission strategy is MRT that satisfies null-shaping constraints given in $Z_k(\boldsymbol{\lambda}_k)$. The MRT strategy is the unique noncooperative strategy of transmitter k , and corresponds to the NE strategy in game theoretical terms [41]. Through the design of the null-shaping constraints in Corollary 1, all Pareto optimal points of the utility region are characterized by transmission strategies that are in NE and satisfy the characterized null-shaping constraints. The interesting observations are as follows. Null-shaping constraints are sufficient to characterize the Pareto boundary of the MISO IC rate region. Moreover, given the null-shaping constraints, the transmitters are required to be noncooperative in order to achieve efficient operating points [40], [2]. This result can be exploited by clever secondary user selection algorithms.

VI. EXTENSION TO MULTIPLE ANTENNAS AT THE RECEIVERS

We discuss the extension to multiple-antennas at the receivers and study which settings can be applied to our framework. Assume that the number of antennas used by transmitter k is N_k and the number of receive antennas at receiver ℓ is R_ℓ . The channel matrix from transmitter k to receiver ℓ is denoted by $\mathbf{H}_{k\ell} \in \mathbb{C}^{R_\ell \times N_k}$. The signal after receive filtering at receiver ℓ is

$$y_\ell = \sum_{k=1}^T \mathbf{z}_\ell^H \mathbf{H}_{k\ell} \mathbf{w}_k s_k + \mathbf{z}_\ell^H \mathbf{n}_\ell, \quad (30)$$

where \mathbf{w}_k and \mathbf{z}_ℓ are transmit and receive beamforming vectors, respectively. Considering a single transmitter k as in Section III, the power gain achieved at receiver ℓ is $x_{k\ell}(\mathbf{w}_k, \mathbf{z}_\ell) = |\mathbf{z}_\ell^H \mathbf{H}_{k\ell} \mathbf{w}_k|^2$. Clearly, if the receive beamforming vectors do not depend on the transmit beamforming vector \mathbf{w}_k , then the framework in this paper can be applied to the setting. For example, the receive beamforming vectors \mathbf{z}_ℓ can be fixed prior to transmission as in singular value decomposition (SVD) based receive beamforming. Let the SVD⁵ of $\mathbf{H}_{k\ell}$ be $\mathbf{U}_{\mathbf{H}_{k\ell}} \mathbf{\Sigma}_{\mathbf{H}_{k\ell}} \mathbf{V}_{\mathbf{H}_{k\ell}}^H$. The SVD based beamforming vector $\mathbf{z}_\ell^{\text{svd}}$ depends only on $\mathbf{H}_{k\ell}$ and is the left singular vector in $\mathbf{U}_{\mathbf{H}_{k\ell}}$ which corresponds to the largest singular value in $\mathbf{\Sigma}_{\mathbf{H}_{k\ell}}$. The framework in this paper can also be applied when the receive beamforming vectors depend only on intended transmitters' beamforming vectors. Assume the intended transmitter of receiver ℓ is k , then the maximum ratio combining (MRC) beamforming vector of receiver ℓ is $\mathbf{z}_\ell^{\text{mrc}}(\mathbf{w}_k) = \mathbf{H}_{k\ell} \mathbf{w}_k$. In this case, the transmit beamforming vectors of the transmitter are not coupled through the receive beamforming vectors. Hence, the power gain region of a single transmitter can be used as a tool to characterize its efficient beamforming vectors. In case one receiver has multiple intended transmitters as receiver 2 in the example in Section II-B. Receiver 2 uses $\mathbf{z}_{2,12} = \mathbf{H}_{(12)2} \mathbf{w}_{12}$ to successively decode the signal from transmitter 12 and uses $\mathbf{z}_{2,2} = \mathbf{H}_{22} \mathbf{w}_2$ to successively decode the signal from transmitter 2. The utility of receiver 2 is $\log_2(1 + |\mathbf{w}_{12}^H \mathbf{H}_{12}^H \mathbf{H}_{12} \mathbf{w}_{12}|^2 + |\mathbf{w}_2^H \mathbf{H}_{22}^H \mathbf{H}_{22} \mathbf{w}_2|^2)$, where the power gains from the transmitters are not coupled by the transmit beamforming vectors.

The framework in this paper will not apply if the receive beamforming vectors depend on several transmit beamforming vectors. An example is the linear minimum mean square error (MMSE) receiver [42]. Assume the intended transmitter of receiver ℓ is k , then the linear MMSE receiver is $\mathbf{z}_\ell^{\text{mmse}}(\mathbf{w}_1, \dots, \mathbf{w}_T) = (\sigma^2 \mathbf{I} + \sum_{j \neq k} \mathbf{H}_{j\ell} \mathbf{w}_j \mathbf{w}_j^H \mathbf{H}_{j\ell}^H)^{-1} \mathbf{H}_{k\ell} \mathbf{w}_k$. Thus the transmit beamforming vectors

⁵ $\mathbf{H} = \mathbf{U}_H \mathbf{\Sigma}_H \mathbf{V}_H^H$ where \mathbf{U}_H is the $R_k \times R_k$ unitary matrix having the left singular vectors, \mathbf{V}_H is the $N_k \times N_k$ unitary matrix having the right singular vectors, and $\mathbf{\Sigma}_H$ is the $R_k \times N_k$ diagonal matrix containing the nonnegative singular values.

$\mathbf{w}_1, \dots, \mathbf{w}_T$ are coupled through the receive beamforming vector. Hence, the power gain-regions of the transmitters are coupled by their beamforming vectors.

VII. CONCLUSIONS

In this work, we consider MISO wireless interference networks in which T transmitters and K receivers share the same spectral band. We provide a framework to determine the transmission strategies that are relevant for Pareto optimal operation. By studying the single transmitter's power gain-region, the properties of efficient transmission are acquired. We prove that the boundary of the gain-region is convex and achieved with single-stream beamforming. Due to the convexity of the boundary of the gain-region, the efficient transmit beamforming vectors can be parameterized by real-valued parameters. We determine and distinguish the conditions under which power control is required. When the number of antennas at the transmitter is greater than or equal to the number of existing receivers, we show that full power transmission achieves all boundary points of the gain region. In this case, the parameterizations of efficient beamforming vectors requires $K - 1$ real-valued parameters between zero and one. When the number of antennas at the transmitter is strictly less than the number of receivers, we characterize the transmission strategies for which power control is required. For this case, an additional real-valued parameter is needed that varies the power level at the transmitter. We apply the single-transmitter framework to the multiple-transmitter case. On determining the important boundary part of each transmitter's gain-region, all Pareto efficient beamforming vectors are characterized. This parameterizations simplifies the design of Pareto efficient resource allocation schemes.

ACKNOWLEDGMENT

The authors would like to thank Yiu Tung Poon, Erik Larsson, Christian Scheunert, Johannes Lindblom, and Eleftherios Karipidis for interesting discussions.

APPENDIX A

PROOF OF LEMMA 1

Note that Ω does not correspond to the joint field of values [32] which is known to be convex and compact. Therefore, we provide a proof of the convexity of Ω and show that it is bounded and closed, thus compact. It is simple to show that Ω is bounded because the power gain at the ℓ th receiver has a finite maximum which is achieved when the transmitter performs MRT to that receiver, i.e., $x_\ell(\mathbf{Q}) \leq x_\ell\left(\frac{\mathbf{h}_\ell \mathbf{h}_\ell^H}{\mathbf{h}_\ell^H \mathbf{h}_\ell}\right) = \|\mathbf{h}_\ell\|^2$. Therefore, the box described by the set $\mathcal{Y} := \left\{ \mathbf{x} : 0 \leq x_\ell \leq \|\mathbf{h}_\ell\|^2 \right\}$, contains Ω ,

i.e., $\Omega \subset \mathcal{Y}$ as illustrated in Fig. 2. The set Ω is closed because the feasible set of transmission strategies, $\mathbf{Q} \succeq 0$ and $\text{tr}(\mathbf{Q}) \leq 1$, is compact and convex. Since every pre-image of a closed set is closed for continuous functions [43], follows that Ω is a closed set.

It remains to prove that Ω is convex. For any two points $\mathbf{x}(\mathbf{Q}_x) \in \Omega$ and $\mathbf{x}(\mathbf{Q}_y) \in \Omega$, we prove that $\mathbf{x}(\mathbf{Q}_z) \in \Omega$, where $\mathbf{x}(\mathbf{Q}_z) = t\mathbf{x}(\mathbf{Q}_x) + (1-t)\mathbf{x}(\mathbf{Q}_y)$ and $t \in [0, 1]$. Any component of $\mathbf{x}(\mathbf{Q}_z)$ is

$$x_\ell(\mathbf{Q}_z) = tx_\ell(\mathbf{Q}_x) + (1-t)x_\ell(\mathbf{Q}_y) = t\mathbf{h}_\ell^H \mathbf{Q}_x \mathbf{h}_\ell + (1-t)\mathbf{h}_\ell^H \mathbf{Q}_y \mathbf{h}_\ell = \mathbf{h}_\ell^H (t\mathbf{Q}_x + (1-t)\mathbf{Q}_y) \mathbf{h}_\ell. \quad (31)$$

Hence, the transmit covariance matrices that achieve the line segment between $\mathbf{x}(\mathbf{Q}_x)$ and $\mathbf{x}(\mathbf{Q}_y)$ are given as $\mathbf{Q}_z(t) = t\mathbf{Q}_x + (1-t)\mathbf{Q}_y$. Accordingly, since \mathbf{Q}_x and \mathbf{Q}_y are positive semidefinite, then $\mathbf{Q}_z(t)$ is positive semidefinite for all $t \in [0, 1]$. In addition, $\mathbf{Q}_z(t)$ fulfills the trace constraint, $\text{tr}(\mathbf{Q}_z(t)) \leq 1$, for all $t \in [0, 1]$ since $\text{tr}(\mathbf{Q}_z) = \text{tr}(t\mathbf{Q}_x + (1-t)\mathbf{Q}_y) = t\text{tr}(\mathbf{Q}_x) + (1-t)\text{tr}(\mathbf{Q}_y) \leq 1$. Therefore, $\mathbf{x}(\mathbf{Q}_z(t))$ also lies in Ω for all $t \in [0, 1]$. Hence, the set Ω is a compact and convex set.

APPENDIX B

PROOF OF LEMMA 2

In order to prove that the transmit covariance matrices from $\hat{\mathcal{S}}$ in (12) achieve points on the boundary of the set Ω in direction \mathbf{e} , $\mathbf{e} \in \mathcal{E}$, we show that for any transmit covariance matrix \mathbf{P} with $\text{tr}(\mathbf{P}) < 1$, a transmit covariance matrix \mathbf{Q} from $\hat{\mathcal{S}}$ can be constructed in which $\mathbf{x}(\mathbf{Q})$ dominates $\mathbf{x}(\mathbf{P})$ in direction \mathbf{e} according to Definition 3. Assume $e_\ell = +1$, we can construct \mathbf{Q} as $\mathbf{Q} = \mathbf{P} + (1 - \text{tr}(\mathbf{P})) \frac{\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp}{\|\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp\|}$, where $\mathbf{Z} = [\mathbf{h}_1, \dots, \mathbf{h}_{\ell-1}, \mathbf{h}_{\ell+1}, \dots, \mathbf{h}_K]$. Clearly, \mathbf{Q} is in the set $\hat{\mathcal{S}}$. Since $K \leq N$, the dimension of the null space of \mathbf{Z} is greater or equal to one, therefore the projection $\mathbf{\Pi}_Z^\perp$ is not equal to the zero vector. The power gain achieved with \mathbf{Q} at the k th receiver, $k \in \mathcal{K}$, is

$$x_k(\mathbf{Q}) = \mathbf{h}_k^H \left(\mathbf{P} + (1 - \text{tr}(\mathbf{P})) \frac{\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp}{\|\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp\|} \right) \mathbf{h}_k. \quad (32)$$

We distinguish two cases. For $k \neq \ell$, the power gain at the k th receiver is

$$\begin{aligned} x_k(\mathbf{Q}) &= \mathbf{h}_k^H \left(\mathbf{P} + (1 - \text{tr}(\mathbf{P})) \frac{\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp}{\|\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp\|} \right) \mathbf{h}_k \\ &= \mathbf{h}_k^H \mathbf{P} \mathbf{h}_k + (1 - \text{tr}(\mathbf{P})) \underbrace{\mathbf{h}_k^H \frac{\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp}{\|\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp\|} \mathbf{h}_k}_{=0} = x_k(\mathbf{P}). \end{aligned} \quad (33)$$

This implies that the gain at receiver $k \neq \ell$ has not changed. For $k = \ell$, the power gain at receiver ℓ is

$$\begin{aligned} x_\ell(\mathbf{Q}) &= \mathbf{h}_\ell^H \left(\mathbf{P} + (1 - \text{tr}(\mathbf{P})) \frac{\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp}{\|\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp\|} \right) \mathbf{h}_\ell = \mathbf{h}_\ell^H \mathbf{P} \mathbf{h}_\ell + (1 - \text{tr}(\mathbf{P})) \mathbf{h}_\ell^H \frac{\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp}{\|\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp\|} \mathbf{h}_\ell \\ &= \mathbf{h}_\ell^H \mathbf{P} \mathbf{h}_\ell + (1 - \text{tr}(\mathbf{P})) \frac{|\mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp \mathbf{h}_\ell|^2}{\|\mathbf{\Pi}_Z^\perp \mathbf{h}_\ell \mathbf{h}_\ell^H \mathbf{\Pi}_Z^\perp\|} > x_\ell(\mathbf{P}). \end{aligned} \quad (34)$$

According to the above results we can construct $\mathbf{Q}, \mathbf{Q} \in \hat{\mathcal{S}}$, such that $x(\mathbf{Q}) \geq^e x(\mathbf{P})$ for any given \mathbf{P} , with $\text{tr}(\mathbf{P}) < 1$. Therefore, the boundary set $\partial^e \Omega, e \in \mathcal{E}$, can be achieved with transmit covariance matrices that fulfill the total power constraint with equality.

APPENDIX C

PROOF OF LEMMA 3

Since the set $\hat{\Omega}$ is convex and compact, the boundary in direction e , $\partial^e \Omega, e \in \mathcal{E}$, can be achieved using the Supporting Hyperplane theorem [29, Theorem 1.5] by the following programming problem

$$\begin{aligned} &\text{maximize} \quad \sum_{\ell=1}^K \lambda_\ell e_\ell \mathbf{h}_\ell^H \mathbf{Q} \mathbf{h}_\ell \\ &\text{subject to} \quad \text{tr}(\mathbf{Q}) = 1, \quad \mathbf{Q} \succeq 0, \end{aligned} \quad (35)$$

where $\boldsymbol{\lambda} \in \mathbf{\Lambda}$ defined in (17). The objective in (35) can be written as

$$\sum_{\ell=1}^K \lambda_\ell e_\ell \mathbf{h}_\ell^H \mathbf{Q} \mathbf{h}_\ell = \sum_{\ell=1}^K \lambda_\ell e_\ell \text{tr}(\mathbf{Q} \mathbf{h}_\ell \mathbf{h}_\ell^H) = \text{tr} \left(\mathbf{Q} \underbrace{\sum_{\ell=1}^K \lambda_\ell e_\ell \mathbf{h}_\ell \mathbf{h}_\ell^H}_{\mathbf{Z}} \right) = \text{tr}(\mathbf{Q} \mathbf{Z}) \leq \sum_{\ell=1}^N \mu_\ell(\mathbf{Q}) \mu_\ell(\mathbf{Z}) \quad (36)$$

where the last inequality holds according to the von Neumann trace inequality of product of matrices [44]. Define $\mathbf{Q} = \mathbf{U}_Q \Sigma_Q \mathbf{U}_Q^H$ and $\mathbf{Z} = \mathbf{U}_Z \Sigma_Z \mathbf{U}_Z^H$ where \mathbf{U}_Q and \mathbf{U}_Z are unitary matrices. Then, the upper bound in (36) can be achieved by choosing $\mathbf{U}_Q = \mathbf{U}_Z$. Hence, the problem in (35) can be equivalently written as

$$\begin{aligned} &\text{maximize} \quad \sum_{\ell=1}^N \mu_\ell(\mathbf{Q}) \mu_\ell(\mathbf{Z}) \\ &\text{subject to} \quad \sum_{\ell=1}^N \mu_\ell(\mathbf{Q}) = 1, \quad \mu_\ell(\mathbf{Q}) \geq 0, \quad \text{for all } \ell = 1, \dots, N. \end{aligned} \quad (37)$$

The solution of this problem is $\mu_N(\mathbf{Q}) = 1$ and $\mu_\ell(\mathbf{Q}) = 0$ if $\ell \neq N$. Thus, \mathbf{Q} is rank-1 and the transmit covariance matrices from $\tilde{\mathcal{S}}$ achieve the boundary of the region $\hat{\Omega}$.

APPENDIX D

PROOF OF THEOREM 1

As in the proof of Lemma 3, the convex boundary of the set $\widehat{\Omega}$ in direction e can be characterized by the solution of the following programming problem

$$\begin{aligned} & \text{maximize} && \sum_{\ell=1}^K \lambda_{\ell} e_{\ell} |\mathbf{w}^H \mathbf{h}_{\ell}|^2 \\ & \text{subject to} && \|\mathbf{w}\|^2 = 1, \end{aligned} \quad (38)$$

The objective function in (38) can be rewritten as

$$y(\mathbf{w}) = \sum_{\ell=1}^K \lambda_{\ell} e_{\ell} |\mathbf{w}^H \mathbf{h}_{\ell}|^2 = \mathbf{w}^H \underbrace{\left(\sum_{\ell=1}^K \lambda_{\ell} e_{\ell} \mathbf{h}_{\ell} \mathbf{h}_{\ell}^H \right)}_{\mathbf{Z}} \mathbf{w}. \quad (39)$$

Note that the matrix \mathbf{Z} in (39) is not necessarily positive semidefinite because the directional vector e can contain negative components. However, it is Hermitian and therefore, the solution to (38) is the eigenvector which corresponds to the largest eigenvalue of \mathbf{Z} .

APPENDIX E

PROOF OF THEOREM 2

The proof is by contradiction. Assume that a power gain tuple $\mathbf{x}_k(\mathbf{w}_k \mathbf{w}_k^H)$ is not on the boundary of the power gain region Ω_k and achieves a Pareto optimal point in the utility region. Since $\mathbf{x}_k(\mathbf{w}_k \mathbf{w}_k^H)$ is not on the boundary of Ω_k , we can find another power gain tuple on the boundary of the power gain region which dominates $\mathbf{x}_k(\mathbf{w}_k \mathbf{w}_k^H)$ in the direction specified in (22). Accordingly, it is possible to increase the utility of at least one receiver without changing the utilities of the other receivers which follows from Assumption 1. Thus, a contradiction is made on the Pareto optimality of $\mathbf{x}_k(\mathbf{w}_k \mathbf{w}_k^H)$. Formally, assume

$$\mathbf{x}_k(\mathbf{w}_k \mathbf{w}_k^H) \notin \partial^{e_k} \Omega_k, \quad (40)$$

and achieves a point on the Pareto boundary of \mathcal{U} , i.e.

$$(u_1(x_{1,1}(\mathbf{w}_1 \mathbf{w}_1^H), \dots, x_{T,1}(\mathbf{w}_T \mathbf{w}_T^H)), \dots, u_K(x_{1,K}(\mathbf{w}_1 \mathbf{w}_1^H), \dots, x_{T,K}(\mathbf{w}_T \mathbf{w}_T^H))) \in \mathcal{PB}, \quad (41)$$

according to Definition 1. By assuming (40), we can find another feasible beamforming vector \mathbf{v}_k such that $\mathbf{v}_k \mathbf{v}_k^H \in \mathcal{S}_k$ and achieves power gain tuple on the boundary of the power-gain region in direction e_k such that

$$x_{k,\ell}(\mathbf{w}_k \mathbf{w}_k^H) e_{k,\ell} \leq x_{k,\ell}(\mathbf{v}_k \mathbf{v}_k^H) e_{k,\ell}, \quad \text{for all } \ell \in \mathcal{K}. \quad (42)$$

Next we distinguish two cases for the inequality in (42) corresponding to $\ell \in \overline{\mathcal{K}}(k)$ or $\ell \in \underline{\mathcal{K}}(k)$.

- 1) Assume the inequality in (42) is strict for $\ell \in \overline{\mathcal{K}}(k)$ with $e_{k,\ell} = +1$ as given in (22), then $x_{k,\ell}(\mathbf{w}_k \mathbf{w}_k^H) < x_{k,\ell}(\mathbf{v}_k \mathbf{v}_k^H)$. The power gains to all other receivers are to stay unchanged such that $x_{k,j}(\mathbf{w}_k \mathbf{w}_k^H) = x_{k,j}(\mathbf{v}_k \mathbf{v}_k^H)$ for $j \neq \ell$. Then,

$$u_\ell(x_{1,\ell}(\mathbf{w}_1 \mathbf{w}_1^H), \dots, x_{T,\ell}(\mathbf{w}_T \mathbf{w}_T^H)) < u_\ell(x_{1,\ell}(\mathbf{w}_1 \mathbf{w}_1^H), \dots, x_{k,\ell}(\mathbf{v}_k \mathbf{v}_k^H), \dots, x_{T,\ell}(\mathbf{w}_T \mathbf{w}_T^H)) \quad (43)$$

holds according to property A in Assumption 1 in section II-A. This result contradicts (41).

- 2) Assume the inequality in (42) is strict for $j \in \underline{\mathcal{K}}(k)$ with $e_{k,j} = -1$. Then, (42) changes to $x_{k,j}(\mathbf{w}_k) > x_{k,j}(\mathbf{v}_k)$. Assuming $x_{k,\ell}(\mathbf{w}_k) = x_{k,\ell}(\mathbf{v}_k)$ for $\ell \neq j$, then

$$u_j(x_{1,j}(\mathbf{w}_1 \mathbf{w}_1^H), \dots, x_{T,j}(\mathbf{w}_T \mathbf{w}_T^H)) < u_j(x_{1,j}(\mathbf{w}_1 \mathbf{w}_1^H), \dots, x_{k,j}(\mathbf{v}_k \mathbf{v}_k^H), \dots, x_{T,j}(\mathbf{w}_T \mathbf{w}_T^H)) \quad (44)$$

holds according to property B in Assumption 1 in section II-A. This result contradicts (41).

Therefore, all points on the Pareto boundary of the utility region are achieved if each transmitter chooses its beamforming vectors to achieve the boundary of the gain-region in the direction specified in (22). These beamforming vectors are characterized in Theorem 1 in Section III-A and for the optimal power control in (20) in Section III-B, which lead to the formulation in Theorem 2.

APPENDIX F

PROOF OF LEMMA 4

The zero-forcing condition on the beamforming vector of transmitter k , denoted as \mathbf{w}_k^{ZF} , $k \in \mathcal{K}$, is

$$x_{k,\ell}(\mathbf{w}_k^{ZF}) : \begin{cases} \geq 0 & k = \ell \\ = 0 & \text{otherwise} \end{cases}. \quad (45)$$

It is possible to fulfill the condition in (45) if $N_k \geq K$ and all channel vectors from transmitter k to all receivers are linearly independent. Here, we give a direct proof. Assume the conditions $\lambda_{k,\ell} = 0$ for $k = \ell$ and $\lambda_{k,\ell} > 0$ for $k \neq \ell$ hold, we have to show that these conditions lead to beamforming vectors that satisfy (45). In order to do this, we have to study the matrix in (21) whose eigenvector, corresponding to the largest eigenvalue, determines the used beamforming vector. Define the matrix \mathbf{M}_k , as

$$\mathbf{M}_k = \sum_{\ell=1}^K \lambda_{k,\ell} e_{k,\ell} \mathbf{h}_{k,\ell} \mathbf{h}_{k,\ell}^H = \underbrace{\lambda_{k,k} \mathbf{h}_{kk} \mathbf{h}_{kk}^H}_{\mathbf{A}_k} - \underbrace{\sum_{\ell \neq k} \lambda_{k,\ell} \mathbf{h}_{k\ell} \mathbf{h}_{k\ell}^H}_{\mathbf{B}_k}, \quad (46)$$

where the direction vector \mathbf{e} is specified as for the MISO IC application. According to the conditions $\lambda_{k,\ell} = 0$ for $k = \ell$ and $\lambda_{k,\ell} > 0$ for $k \neq \ell$, \mathbf{A}_k in (46) is equal to zero, and hence \mathbf{M}_k is

negative semidefinite. The largest eigenvalue of \mathbf{M}_k is therefore zero since $N_k \geq K$, and we can write $(\sum_{\ell \neq k} \lambda_{k,\ell} \mathbf{h}_{k\ell} \mathbf{h}_{k\ell}^H) \mathbf{w}_k = \mathbf{0}$. Since all channel vectors are linearly independent, the eigenvector associated with the largest eigenvalue produces zero gain on any of the interference channel vectors. It is clearly seen that if $\lambda_\ell = 0$ for any $\ell \neq k$, then the largest eigenvector does not necessarily produce zero gain on this receiver.

APPENDIX G

PROOF OF COROLLARY 1

We prove that the gains achieved by the beamforming vectors in (29) are equal to the gains achieved by the beamforming vectors given in (21). Define the matrix \mathbf{M}_k , with the direction vector $\mathbf{e}_{k,\ell}$ given in (22), as

$$\mathbf{M}_k = \sum_{\ell=1}^K \lambda_{k,\ell} \mathbf{e}_{k,\ell} \mathbf{h}_{k,\ell} \mathbf{h}_{k,\ell}^H = \underbrace{\sum_{\ell \in \overline{\mathcal{K}}(k)} \lambda_{k,\ell} \mathbf{h}_{k\ell} \mathbf{h}_{k\ell}^H}_{\mathbf{A}_k} + \underbrace{\sum_{\ell \in \underline{\mathcal{K}}(k)} -\lambda_{k,\ell} \mathbf{h}_{k\ell} \mathbf{h}_{k\ell}^H}_{\mathbf{B}_k} \quad (47)$$

The matrices \mathbf{M}_k , \mathbf{A}_k and \mathbf{B}_k are Hermitian matrices of size $N_k \times N_k$. The eigenvalues of \mathbf{M}_k are real and we always consider them ordered in nondecreasing order, i.e., $\mu_1(\mathbf{M}_k) \leq \mu_2(\mathbf{M}_k) \leq \dots \leq \mu_{N_k}(\mathbf{M}_k)$. \mathbf{A}_k consists of sum of positive semidefinite matrices. Hence, $\mathbf{A}_k \succeq 0$ and $\text{rank}(\mathbf{A}_k) \leq |\overline{\mathcal{K}}(k)|$, i.e.,

$$0 \leq \mu_{|\overline{\mathcal{K}}(k)|+1}(\mathbf{A}_k) \leq \dots \leq \mu_{N_k}(\mathbf{A}_k), \text{ and } \mu_1(\mathbf{A}_k) = \dots = \mu_{N_k-|\overline{\mathcal{K}}(k)|}(\mathbf{A}_k) = 0. \quad (48)$$

\mathbf{B}_k consists of the sum of the negative of positive semidefinite matrices. Hence, $\mathbf{B}_k \preceq 0$ and $\text{rank}(\mathbf{B}_k) \leq |\underline{\mathcal{K}}(k)|$, which leads to the following properties on the eigenvalues:

$$\mu_1(\mathbf{B}_k) \leq \dots \leq \mu_{|\underline{\mathcal{K}}(k)|}(\mathbf{B}_k) \leq 0, \text{ and } \mu_{|\underline{\mathcal{K}}(k)|+1}(\mathbf{B}_k) = \dots = \mu_{N_k}(\mathbf{B}_k) = 0. \quad (49)$$

Next, we study the eigenvalues of $\mathbf{M}_k = \mathbf{A}_k + \mathbf{B}_k$. According to Weyl's inequality of the eigenvalues of the sum of Hermitian matrices [36, Theorem 4.3.7] the following properties are gained:

$$\mu_{N_k-|\overline{\mathcal{K}}(k)|}(\mathbf{M}_k) \leq \mu_{N_k-|\overline{\mathcal{K}}(k)|}(\mathbf{A}_k) + \mu_{N_k}(\mathbf{B}_k) = 0 \quad (50)$$

$$\mu_{|\underline{\mathcal{K}}(k)|+1}(\mathbf{M}_k) \geq \mu_1(\mathbf{A}_k) + \mu_{|\underline{\mathcal{K}}(k)|+1}(\mathbf{B}_k) = 0 \quad (51)$$

The eigenvalues of \mathbf{M}_k are ordered in nondecreasing order. Therefore, the following eigenvalues of \mathbf{M}_k are always equal to zero: $\mu_{|\underline{\mathcal{K}}(k)|+1}(\mathbf{M}_k) = \dots = \mu_{N_k-|\overline{\mathcal{K}}(k)|}(\mathbf{M}_k) = 0$. In addition, the smallest $|\underline{\mathcal{K}}(k)|$ eigenvalues of \mathbf{M}_k are nonpositive.

If the dimension of space is larger than the number of receivers, i.e., $N_k \geq |\mathcal{K}|$, then there would be at least $N_k - |\underline{\mathcal{K}}(k)| - |\overline{\mathcal{K}}(k)|$ eigenvalues of \mathbf{M}_k that are zero. For the eigenvectors corresponding to those eigenvalues, the eigenvalue equation is written as $(\sum_{\ell \in \overline{\mathcal{K}}(k)} \lambda_{k,\ell} \mathbf{h}_{k\ell} \mathbf{h}_{k\ell}^H + \sum_{\ell \in \underline{\mathcal{K}}(k)} -\lambda_{k,\ell} \mathbf{h}_{k\ell} \mathbf{h}_{k\ell}^H) \mathbf{v}_i = \mathbf{0}$, for all $i = |\underline{\mathcal{K}}(k)| + 1, \dots, N_k - |\overline{\mathcal{K}}(k)|$. Then, for all $\ell \in \mathcal{K}$,

$$(\lambda_{k,\ell} \mathbf{h}_{k\ell} \mathbf{h}_{k\ell}^H) \mathbf{v}_i = 0, \quad \text{for all } i = |\underline{\mathcal{K}}(k)| + 1, \dots, N_k - |\overline{\mathcal{K}}(k)|. \quad (52)$$

The set of eigenvectors of \mathbf{M}_k , $\{\mathbf{v}_1, \dots, \mathbf{v}_{N_k}\}$, form an orthonormal set, i.e. $\|\mathbf{v}_i\| = 1$ for all $i = 1, \dots, N_k$ and $\mathbf{v}_i^H \mathbf{v}_j = 0$ for $i \neq j$. Therefore, we can write $\sum_{\ell=1}^{N_k} \mathbf{v}_\ell \mathbf{v}_\ell^H = \mathbf{I}$, which gives

$$\mathbf{v}_{N_k}(\boldsymbol{\lambda}) \mathbf{v}_{N_k}^H(\boldsymbol{\lambda}) = \mathbf{I} - \sum_{\ell=1}^{N_k-1} \mathbf{v}_\ell(\boldsymbol{\lambda}) \mathbf{v}_\ell^H(\boldsymbol{\lambda}) = \mathbf{I} - G_k(\boldsymbol{\lambda}) G_k^H(\boldsymbol{\lambda}) = \boldsymbol{\Pi}_{G_k(\boldsymbol{\lambda})}^\perp, \quad (53)$$

where $G_k(\boldsymbol{\lambda}) = [\mathbf{v}_1(\boldsymbol{\lambda}), \dots, \mathbf{v}_{N_k-1}(\boldsymbol{\lambda})]$. Let the matrix $Z_k(\boldsymbol{\lambda})$ consist of the eigenvectors of $G_k(\boldsymbol{\lambda})$ excluding the eigenvectors that satisfy (52), i.e.,

$$Z_k(\boldsymbol{\lambda}) = [\mathbf{v}_1(\boldsymbol{\lambda}), \dots, \mathbf{v}_{|\underline{\mathcal{K}}(k)|}(\boldsymbol{\lambda}), \mathbf{v}_{N_k-|\overline{\mathcal{K}}(k)|+1}(\boldsymbol{\lambda}), \dots, \mathbf{v}_{N_k-1}(\boldsymbol{\lambda})], \quad (54)$$

then for any $\mathbf{g} \in \mathbb{C}^{N_k}$ we can write

$$\left| \mathbf{g}^H \frac{\boldsymbol{\Pi}_{Z_k(\boldsymbol{\lambda})}^\perp \mathbf{h}_{k\ell}}{\|\boldsymbol{\Pi}_{Z_k(\boldsymbol{\lambda})}^\perp \mathbf{h}_{k\ell}\|} \right|^2 = \left| \mathbf{g}^H \frac{\boldsymbol{\Pi}_{G_k(\boldsymbol{\lambda})}^\perp \mathbf{h}_{k\ell}}{\|\boldsymbol{\Pi}_{G_k(\boldsymbol{\lambda})}^\perp \mathbf{h}_{k\ell}\|} \right|^2 = \left| \mathbf{g}^H \frac{\mathbf{v}_{N_k}(\boldsymbol{\lambda}) \mathbf{v}_{N_k}^H(\boldsymbol{\lambda}) \mathbf{h}_{k\ell}}{\|\mathbf{v}_{N_k}(\boldsymbol{\lambda}) \mathbf{v}_{N_k}^H(\boldsymbol{\lambda}) \mathbf{h}_{k\ell}\|} \right|^2 = |\mathbf{g}^H \mathbf{v}_{N_k}(\boldsymbol{\lambda})|^2, \quad (55)$$

where $\ell \in \mathcal{K}$. Hence, the same power gains are achieved with $\frac{\boldsymbol{\Pi}_{Z_k(\boldsymbol{\lambda})}^\perp \mathbf{h}_{k\ell}}{\|\boldsymbol{\Pi}_{Z_k(\boldsymbol{\lambda})}^\perp \mathbf{h}_{k\ell}\|}$ as with $\mathbf{v}_{N_k}(\boldsymbol{\lambda})$.

REFERENCES

- [1] E. A. Jorswieck, "Beamforming in interference networks: Multicast, MISO IFC and secrecy capacity," in *Proc. IZS*, 2010, invited.
- [2] E. A. Jorswieck and R. Mochaourab, "Beamforming in underlay cognitive radio: Null-shaping design for efficient Nash equilibrium," in *Proc. of International Workshop on Cognitive Information Processing (CIP)*, Jun. 2010, pp. 476–481, invited.
- [3] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "On the complexity of optimal coordinated downlink beamforming," in *Proc. ICASSP*, 2010.
- [4] E. A. Jorswieck and E. G. Larsson, "The MISO interference channel from a game-theoretic perspective: A combination of selfishness and altruism achieves pareto optimality," in *Proc. ICASSP*, Apr. 2008, pp. 5364–5367.
- [5] —, "Monotonic optimization framework for the MISO interference channel," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2159–2168, Jul. 2010.
- [6] R. Zakhour and D. Gesbert, "Coordination on the MISO interference channel using the virtual sinr framework," in *Proc. WSA*, 2009.
- [7] R. Zakhour, K. M. Z. Ho, and D. Gesbert, "Distributed beamforming coordination in multicell MIMO channels," in *Proc. VTC*, Apr. 2009.

- [8] Z. K. M. Ho, , and D. Gesbert, "Spectrum sharing in multiple-antenna channels: A distributed cooperative game theoretic approach," in *Proc. PIMRC*, Sept. 2008, pp. 1–5.
- [9] J. Lindblom and E. Karipidis, "Cooperative beamforming for the MISO interference channel," in *Proc. of 16th European Wireless Conference (EW'10)*, 2010.
- [10] Z. K. M. Ho and D. Gesbert, "Balancing egoism and altruism on the interference channel: The MIMO case," in *Proc. ICC*, May. 2010, pp. 1–5.
- [11] R. Mochaourab, E. A. Jorswieck, K. M. Z. Ho, and D. Gesbert, "Bargaining and beamforming in interference channels," in *Proc. Asilomar*, 2010, to appear.
- [12] Y. Lebrun, J. Nsenga, V. Ramon, A. Bourdoux, F. Horlin, and R. Lauwereins, "Maximum SINR-based beamforming for the MISO interference channel," in *Proc. of 17th European Signal Processing Conference*, Aug. 2009.
- [13] E. A. Jorswieck, E. G. Larsson, and D. Danev, "Complete characterization of the pareto boundary for the MISO interference channel," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5292–5296, Oct. 2008.
- [14] X. Shang, B. Chen, and H. V. Poor, "On the optimality of beamforming for multi-user MISO interference channels with single-user detection," *CoRR*, vol. abs/0908.3171, 2009.
- [15] R. Zhang and S. Cui, "Cooperative interference management with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5450–5458, Oct. 2010.
- [16] A. Goldsmith, S. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.
- [17] E. Bjornson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4298–4310, Aug. 2010.
- [18] E. Bjornson, M. Bengtsson, and B. Ottersten, "Pareto characterization of the multicell MIMO performance region with simple receivers," submitted to ICC 2011.
- [19] C.-B. Chae, I. Hwang, J. Robert W. Heath, and V. Tarokh, "Interference aware-coordinated beamform system in a two-cell environment," *IEEE J. Sel. Areas Commun.*, 2010, to appear.
- [20] A. Gershman, N. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 62–75, May 2010.
- [21] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.
- [22] D. Tomecki and S. Stanczak, "On feasible SNR region for multicast downlink channel: Two user case," in *Proc. ICASSP*, Mar. 2010, pp. 3474–3477.
- [23] J. Lindblom, E. Karipidis, and E. G. Larsson, "Selfishness and altruism on the MISO interference channel: The case of partial transmitter CSI," *IEEE Commun. Lett.*, vol. 13, no. 9, pp. 667–669, 2009.
- [24] J. Lindblom, E. G. Larsson, and E. A. Jorswieck, "Parameterization of the MISO IFC rate region: the case of partial channel state information," *IEEE Trans. Commun.*, vol. 9, no. 2, pp. 500–504, Feb. 2010.
- [25] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," in *Proc. ISIT*, 2004, p. 174.
- [26] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge: Cambridge Univ. Press, 2005.
- [27] M. Luby, "LT codes," in *Proc. FOCS*, 2002, pp. 271–280.
- [28] D. J. C. MacKay, "Fountain codes," *Communications, IEE Proceedings*, vol. 152, no. 6, pp. 1062–1068, Dec. 2005.

- [29] H. Tuy, *Convex Analysis and Global Optimization*. Kluwer Academic Publishers, 1998.
- [30] —, “Monotonic optimization: Problems and solution approaches,” *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000.
- [31] E. G. Larsson and E. A. Jorswieck, “Competition versus cooperation on the MISO interference channel,” *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1059–1069, Sept. 2008.
- [32] G. P. Barker, “Applications of the joint angular field of values,” *Proc. Amer. Math. Soc.*, vol. 91, no. 3, pp. 331–335, 1984.
- [33] F. F. Bonsall and J. Duncan, *Numerical Ranges. II*. Cambridge: Cambridge Univ. Press, 1973.
- [34] C.-K. Li and Y.-T. Poon, “Convexity of the joint numerical range,” *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 2, pp. 668–678, 2000.
- [35] X. Shang, B. Chen, and H. V. Poor, “Multi-user MISO interference channels with single-user detection: Optimality of beamforming and the achievable rate region,” *CoRR*, vol. abs/0907.0505, 2009.
- [36] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge Univ. Press, 1985.
- [37] S. Katz, A. Tal, and R. Basri, “Direct visibility of point sets,” *ACM Trans. Graph.*, vol. 26, no. 3, p. 24, 2007.
- [38] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [39] G. Scutari, D. Palomar, J.-S. Pang, and F. Facchinei, “Flexible design of cognitive radio wireless systems,” *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 107–123, Sept. 2009.
- [40] E. A. Jorswieck and R. Mochaourab, “Beamforming in underlay cognitive radio: Null-shaping constraints and greedy user selection,” in *Proc. CROWNCOM*, Jun. 2010, pp. 1–5, invited.
- [41] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. The MIT Press, 1994.
- [42] P. Viswanath, V. Anantharam, and D. Tse, “Optimal sequences, power control, and user capacity of synchronous CDMA systems with linear MMSE multiuser receivers,” *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 1968–1983, Sep. 1999.
- [43] W. Rudin, *Analysis*, 2nd ed. R. Oldenbourg Verlag, 2002.
- [44] L. Mirsky, “On the trace of matrix products,” *Mathematische Nachrichten*, vol. 20, pp. 171–174, 1959.